ELSEVIER

FEBS Letters 338 (1994) 267–271

FEBS 13593

FEBS LETTERS

The atomic structure of Carnation Mottle Virus capsid protein

E.Yu. Morgunova[a], Z. Dauter[b], E. Fry[c], D.I. Stuart[c], V.Ya. Stel'mashchuk[a], A.M. Mikhailov[a,*], K.S. Wilson[b], B.K. Vainshtein[a]

[a]*Institute of Crystallography, Russian Academy of Sciences, Leninsky Pr. 59, Moscow 117333, Russian Federation*
[b]*European Molecular Biology Laboratory, c/o DESY, Notkestrasse 85, D-2000 Hamburg 52, Germany*
[c]*Laboratory of Molecular Biophysics, Oxford University, The Rex Richards Building, South Parks Road, Oxford OX1 3QU, UK*

Received 8 December 1993

Abstract

The structure of the Carnation Mottle Virus (CMtV) capsid protein has been determinated at 3.2 Å resolution by the method of molecular replacement. Three-dimensional data were collected from a small number of crystals (sp.g. I23, a = 382.6 Å) using the synchrotron radiation with an image plate as detector. The coordinates of Tomato Bushy Stunt Virus (TBSV) were used as a searching model. Refinement of the coordinates of 7,479 non-hydrogen atoms performed by the program XPLOR, has led to an R-factor of 18.3%. It was found that the amino acid chain fold of capsid protein is very similar to that in other icosahedral viruses. However, there are some differences in the contact regions between protein subunits and also the lack of the β-annulus around the 3-fold icosahedral axes. The structural and biochemical results lead us to consider an alternative assembly pathway.

Key words: Carnation Mottle Virus; Capsid protein; X-ray analysis; Assembly mechanism

1. Introduction

Carnation Mottle Virus is a typical representative of Carmovirus group [1]. Isometric viruses of this group have a host range among different plants, and CMtV is widespread in commercial stocks of carnation.

The virus particle consists of the protein shell, a molecule of genome RNA and two short non-genome RNA fragments. The capsid is composed of 180 copies of the same protein with molecular weight of 37,787 Da arranged in accordance with T = 3 icosahedral symmetry. The amino acid sequence of the coat protein determined from the cloned cDNA nucleotide sequence consists of 347 amino acid residues [2].

Although CMtV belongs to a group of viruses with properties different from, for example, the Tombusvirus group, its viral envelope was found to possess a tertiary and quaternary structure very similar to that of the TBSV structure previously determined [3,4]. It was proposed that the structure of the CMtV coat protein may also be based on the same folding and assembly principles.

Carrington et al. [4] reported sequence alignments between coat proteins of TBSV, Turnip Crinkle Virus (TCV), Southern Bean Mosaic Virus (SBMV) and CMtV. It has been shown that the arms and S-domains of these viruses have only 18 positions at which all four sequences coincide. Although there are many other positions with conservative substitutions, the conserved residues do not appear to have any special role in stabilizing of β-barrel fold or in mediating subunit interaction. They are found both in strands and in loops; their side-chains face both the interior and the surface of the domain. The secondary structure elements are invariant in position and length in the four viruses. The loop between βC and αA is six residues longer in TBSV than in TCV and CMtV, but in SBMV this loop is shorter than in the other three viruses. In the turn between βF and βG in SBMV there is a large external insertion absent in TBSV. In CMtV and TCV it is 1–2 residues shorter.

In this paper we describe the crystallographic study of the three-dimensional structure of CMtV coat protein at 3.2 Å resolution and the results of disassembly–assembly biochemical experiments on the disassembly–assembly of virus particles from a mixture of coat protein and RNA.

2. Materials and methods

2.1. Crystallization

CMtV was isolated from *Dianthus caryophyllus* L. by the PEG precipitation technique and purified using a sucrose gradient [5].

The pure CMtV suspension was crystallized at room temperature by sitting drop vapor diffusion using ammonium sulphate as a precipitant. The crystals most suitable for X-ray analysis were obtained in 40 μl droplets of 0.1 M Tris-HCl buffer solution containing 40–50 mg/ml of

*Corresponding author. Fax: (7) (095) 135 10 11.

SSDI 0014-5793(93)E1510-S

virus and 10% saturated ammonium sulphate. The equilibrating solution consisted of 0.1 M Tris-maleic (mal)/NaOH, pH 5.03 with 25% saturated ammonium sulphate. Either 1.7-heptandiol or PEG 300 were added to lessen the number of pellets. The crystals reached the size of 0.7–0.8 mm within 2–3 weeks.

These crystals belong to cubic space group I23, a = 382 Å, z = 2.

2.2. Data collection

X-ray data to 3.2 Å resolution were collected on the EMBL X11 beam line during the parasitic mode of operation of DORIS storage ring at 5.303 GeV (EMBL, c/o DESY, Hamburg, Germany). The wavelength for CMtV crystals was set to 0.96 to avoid rapid radiation damage add minimize the effects of absorption. The detector was an imaging plate scanner with a plate radius of 110 mm, developed at EMBL by J. Hendrix and A. Lentfer (unpublished). Only four crystals and 12 h on beam-line were necessary to obtain the whole data set. The results of the data collection are presented in Table 1. Data were integrated using the program DENZO (Z. Otwinowski).

2.3. Structure determination

The three-dimensional structure of CMtV was determined by the molecular replacement method with the molecule of coat protein of TBSV serving as a starting model at 2.9 Å resolution. The first phases set was obtained at 6 Å resolution. These phases were improved by iterative averaging [6] and solvent flattening using the automatic envelope determination procedure and protocol developed for the structure determination of Foot-and-Mouth Disease Virus (to maximize the phasing power only with five-fold non-crystallographic redundancy) [7]. Phase-extension to 3.2 Å proceeded in small steps (180 cycles) with repeated envelope determinations. The electron density map at 3.5 Å resolution allowed to identify about 80% of amino acid residues in the three quasi-equivalent subunits of coat protein. The building of the amino acids residues into this map was made using FRODO [8]. All calculations for structure determination and refinement have been carried out at the Laboratory of Molecular Biology, Oxford University.

2.4. Refinement

Refinement has been performed using XPLOR version 2.1 (Jack-Levitt refinement) [9] (140,248 observations and 7,479 independent non-hydrogen atoms) with 5-fold non-crystallographic symmetry constraints. A bulk solvent correction was used with the contribution of partial structure factors weighted by a scale factor of 0.36 and B-factor of 90 (as determined to be optimal). An R-factor of 35.4% for the initial model in the resolution range 6.0–3.5 Å fell down to 18.3% for the final model of CMtV, containing 7,479 atoms, three calcium ions and three sulphate ions. No water molecules are included in the model. The standard deviations for restrained parameters of the last refinement cycle and the R-factor changes are presented in Table 2.

2.5. Virus dissociation and reassembly

The virus dissociation was performed by the method of 'jump' increasing pH. Ten μl of virus suspension with concentration 33 mg/ml in 0.1 M Tris-HCl buffer, pH 7, were added to 3 ml of 0.1 M glycine buffer, pH 10.32. The complete dissociation of virus particles has been observed after 24 h.

The CMtV coat protein was prepared by the method described by Rossmann and Ericson for SBMV [10] with 6 M LiCl, EDTA in 0.1 M Tris-HCl buffer pH 6.9.

Assembly experiments were carried out by the dialysis of mixture of coat protein and RNA in 0.1 M Glycine-NaOH buffer, pH 10.32, against 0.1 M Tris-HCl/NaOH, pH 6–7, or 0.1 M Na-acetate buffer, pH 5.22, at room temperature.

Samples for electron microscopy were negatively strained with 2% uranyl acetate. Electron micrographs were taken on a Philips 400 at an instrument magnification of 50,000 × and an accelerating voltage of 80 kV.

3. Results and discussion

The electron density map calculated with Fobs and refined phases (Fig. 1a,b) was extremely clear and could be interpreted unambiguously in terms of the atomic model with excellent stereochemistry.

As was expected, the overall architecture of the CMtV capsid is similar to that described for the other plant viruses and is composed of 180 chemically identical protein subunits arranged in a T = 3 surface lattice, i.e. with 3 subunits, referred A,B,C in icosahedral asymmetric unit (Fig. 2a). The comparison of the A, B and C subunits showed that the packing of the chemically identical protein molecules in the virus capsid could change both conformation and arrangement of the secondary structure elements turning these subunits into the quasi-equivalent ones.

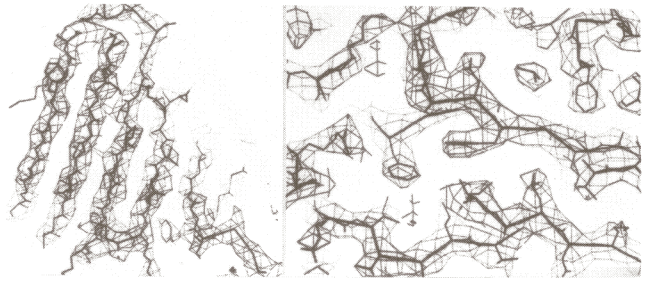
Each protein subunit (Fig. 2b,c) consists of two main ordered domains: the S-surface domain (170 residues, Ser^{82}–Lys^{244}) and the P-domain, forming a protrusion into the solvent (Gln^{250}–Met^{348} residues). These domains are connected by flexible a hinge h (Thr^{245}–Ser^{249}). In contrast to the TBSV protein structure as well as to all other known coat protein structures of icosahedral viruses, the electron density maps did not show any ordered density for the N-terminal arm of any of three protein subunits.

The external P domain is less well ordered than the S domain (mean B-values, which are measures of mobility, are 42.2 $Å^2$ for P compared to 22.1 $Å^2$ for S). The secondary structure of the P-domain is represented by two antiparallel β-sheets (Table 3). The bigger β-sheet, composed

Table 1
Details of the data collection

Crystal	Resolution (Å)	Number of unique reflections	R_{merge} (%)	Completeness (%)
CMVL	5.5	30,716	6.0	97.8
CMV1	3.2	59,258	7.0	52.8
CMV6	3.2	74,913	7.5	66.8
CMV7	3.2	70,274	8.8	62.8
whole data set	6–3.2	140,483	8.2	91.0

$R_{merge} = \Sigma\Sigma I_{h,i} - < I >_h / \Sigma < I >_h$, where $< I >_h = \Sigma I_{h,i} / N_h$.
N_h is the number of measured equivalents.
$I_{h,i}$ is the measured intensity, and i = 1, N_h.



(a) (b)

Fig. 1. The portions of the electron density map after non-crystallographic symmetry averaging at a resolution of 3.5 Å. The map shown was calculated with phases derived from the phase extension procedure and is contoured at an arbitrary level. (a) β-strands in the small β-sheets of the P domain; (b) a portion of the β-strand of the S domain.

of 5 strands, is placed near 2-fold axes and interacts with the smaller 4-stranded sheet of the adjacent subunit. It was found that the strongest interactions are dimer associations driven by P-domain contacts between (i) the A and B subunits clustered around the icosahedral 5-fold axes and (ii) the C subunits related by icosahedral 2-fold axes.

The major structural pattern of S-domain is an 8-stranded antiparallel β-barrel consisting of two β sheets one of which is twisted. Along with the β-strands the S-domain has two helical regions arranged as in the TBSV protein [3].

The structures of both P and S domains are very similar to those observed in TBSV coat protein. In the light of our data, we now know that 33% of the residues in the S domain are identical in sequence and occupy similar positions in both CMtV and TBSV. The corresponding figure for the P domain is only 14%. The more detailed comparison shows that the protein subunits in CMtV are some 39 residues shorter. A structural alignment of the subunits of these viruses shows this shortening to occur by the deletion of 21 residues from the N-terminus, 8 residues from the C-terminus, 3 residues from the S domain and 7 residues from the P-domain.

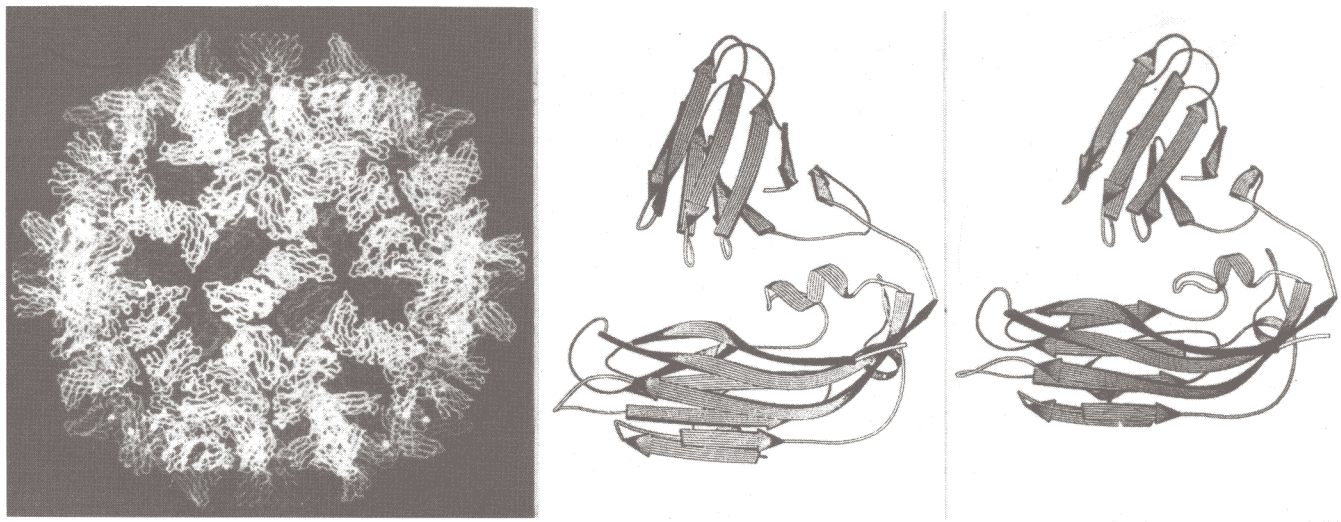
There are two conformations of the hinge for the A/B and C/C interaction region. One hinge conformation (termed 'down') is found in subunits A and B and another one in C-subunits ('up' conformation). In CMtV protein the hinge angle, defined as the angle between the long axis of the P-domain and the 2-fold axis, is approximately 20° smaller in C-subunit dimers than in A/B dimers. This is also the case in TBSV, however both CMtV hinge angles are approximately 10° smaller than in TBSV.

Very strong peak is clearly visible on the electron density map in a region of contacts between C and B protein subunits (Fig. 3a). This density is thought to be with calcium ions coordinated by the residues Asp^{158}, Asp^{161} of the B subunit and Asp^{200}, Asn^{128} of the C-subunit. Harrison suggested the presence of calcium ions between every pair of protein subunits (A/B, A/C and C/B) in CMtV as in TBSV in which there are 6 quasi-equivalent calcium ion binding sites within each icosahedral asymmetric unit [3,4]. However, our results show that only one of the hypothetical Ca-position is occupied despite the proposal made on the basis of sequence alignments. In subunits A and B, a loop from the P domain, which in subunit C and the TBSV coat protein contribute an extra strand to the β-sheet, is pulled down, effectively destroying the calcium binding environment (Fig. 3b).

Table 2
Summary of the refinement statistics

Procedure	Resolution (Å)	R-factor (%)	r.m.s. deviation from ideality Δ_{bonds}, Δ_{angles}, $\Delta_{B\text{-}factors}$ (bonds and angles)
Starting model	3.5	35.4	
Prepare stage (50 steps of J-L)	3.5	27.2	
Prepare stage (200 steps of J-L)	3.5	21.1	
B-factor refinement (20 steps)	3.5	17.8	0.036; 5.223; 2.365; 3.859
Prepare stage (40 steps of J-L)	3.5	19.0	
B-factor refinement (10 steps)	3.5	17.4	
Prepare stage (30 steps of J-L)	3.5	17.3	0.019; 4.2; 4.9; 7.1
Prepare stage (70 steps of J-L)	3.2	18.3	0.021; 4.15; 5.0; 7.0

R-factor = $\Sigma_h (F_{h,obs} - F_{h,calc})/\Sigma_h F_{h,obs}$.
F_{obs} is the observed structure amplitude of reflection h.
F_{calc} is the calculated structure amplitude of atom h.



(a) (b) (c)

Fig. 2. (a) A Cα trace of the CMtV capsid. The subunits parking around the 5-fold axes called the A-subunits, B-subunits are near the quasi-2 axes, and C subunits are those linked by the icosahedral 2-fold axes. The calcium binding site between the B and C subunits is marked by a white sphere and also the sulphate ions are marked like a tetrahedral ball model. (b) A ribbon model of the CMtV capsid protein A/B subunit. (c) C-subunit.

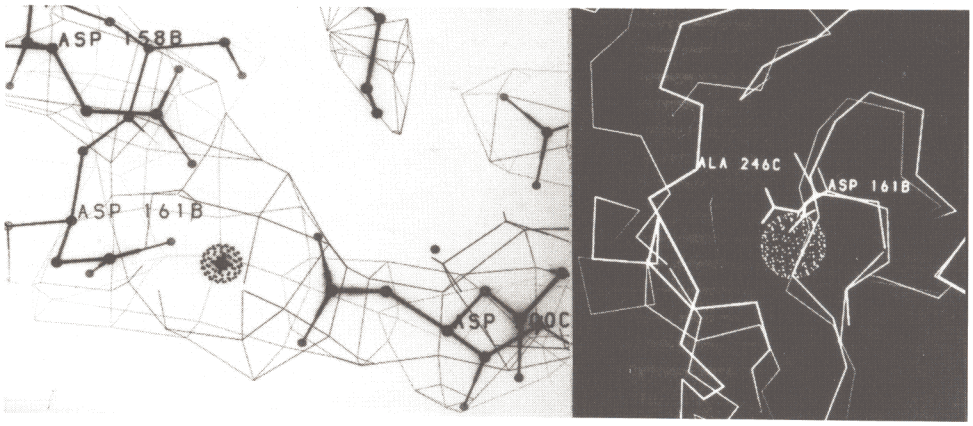
In the investigation of the TBSV coat structure it was shown that the calcium ions between the quasi-equivalent subunits are one of two important elements for the control of assembly and disassembly of virus particles. The second one of these elements is the ordered polypeptide 'arms', preceding the S domain in the primary structure of the subunit that appear to provide a T = 1 scaffold defining the dimensions of the T = 3 particles. In TBSV this arm extends along the 2-fold axis and intertwines with a neighboring C-terminal arm making a strong knot near the 3-fold axis. The A/B dimers are added into this association. The calcium ions play the trigger role in this mechanism of assembly [11,12]. It was quite unexpected that C-subunits of CMtV coat protein would lack ordered N-terminal arms.

The results of the biochemical experiments for dissociation and reassembly of virus particles showed that the pH, ionic strength and the salt concentration can be varied in a wide range and this process is readily reversed (Fig. 4). Apparently, such mobility of the virus shell can be explained by the presence of the large hole around the 3-fold axes at the position of β-annulus in TBSV and the weak contacts near the quasi-3-fold axes.

The yield of native-like (T = 3) particles in the virus

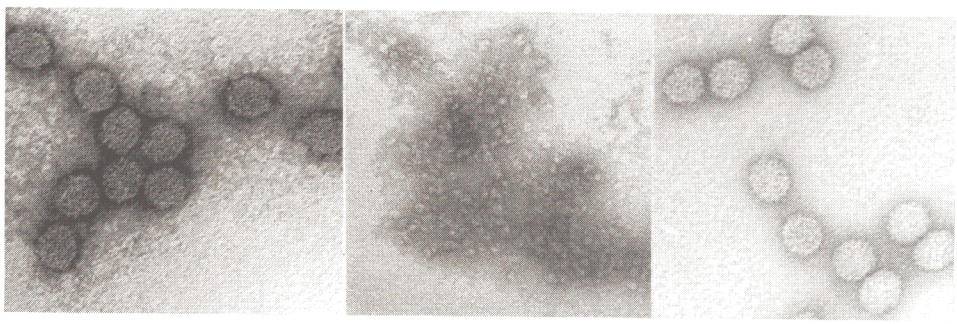
Table 3
The amino acid sequence with the elements of the secondary structure which was determined by the algorithm of Kabsch and Sander [13] using the program DSSP

```
82                  90                  100                 110
S M T M S K T E L L S T V K G T T G V I P S F E D W V V S P R N V
              βB                                  βC
120                 130                 140
A V F P Q L S L L A T N F N K Y R I T A L T V K Y S P A C S F E T N
              αA                  βD
150                 160                 170                 180
G R V A L G F N D D A S D T P P T T K V G F Y D L G K H V E T A A
          βE                                  αB                  βF
190                 200                 210
Q T A K D L V I P V D G K T R F I R D S A S D D A K L V D F G R I V
          βG                                  αC                  βH
220                 230                 240
L S T Y G F D K A D T V V G E L F I Q Y T I V L S D P T K T A K
                                    βI
250                 260                 270                 280
I S Q A S N D K V S D G P T Y V V P S V N G N E L Q L R V V A A G K

290                 300                 310
W C I I V R G T V E G G F T K P T L I G P G I S G D V D Y E S A R P

320                 330                 340
I A V C E L V T Q M E G Q I L K I T K T S A E Q P L Q W V V Y R M
```



(a) (b)

Fig. 3. (a) A portion of the electron density map in the region between B and C subunits, which was interpreted as a calcium-binding site. (b) A superimposition of the calcium binding site between B and C subunits and a 'quasi-equivalent' site is shown between A and B. The residue Ala^{246C} marks the hinge region between S and P domains. Only one of the four calcium ion ligands – Asp^{161B} adopts a different conformation in the quasi-equivalent site.



(a) (b) (c)

Fig. 4. The dissociation and reassembly of the CMtV particle. (a) Electron micrograph of the native virus particles. (b) The mixture of the protein subunits and virus RNA before dialysis. (c) Virus particles, reassembled by 24 h dialysis against Na-acetate buffer, pH 5.22. Magnification: 200,000 ×.

reassembly was typically greater than 80%. Moreover, we could see the other products of reassembly – ellipsoidal particles with the intermediate size between T = 1 and T = 3 shells, and the particles with a greater size.

Our structural and biochemical results allowed for the proposal that the morphological unit of the CMtV particle assembly is a dimer and that there is no difference between A/B and C/C dimers in solution. The self-assembly starts from the organization of complex RNA with the minor protein. The addition of the dimers to the initial intermediate is realized by the charge interactions in which and calcium ions take part. At the same time the calcium ions could be the trigger for the expansion and dissociation of the virus particle which changes the hinge conformation in the C-subunits.

The appearance of the β-annulus structures in the other viruses as a result of the virus evolution makes the virus shell more stable and possibly changes some other processes, for instance, the procedure of the interaction with the host cell.

References

[1] Carrington, J.C. and Morris, T.J. (1988) in: Plant Viruses (R. Koening Ed.) p. 73, Plenum Press, New York.
[2] Gulley, H., Carrington, J.C., Balazs, E., Jonard, G., Richards, K. and Morris, T.J. (1985) Nucleic Acids Res. 13, 6663–6677.
[3] Olson, A.J., Bricogne, G. and Harrison, S.C. (1983) J. Mol. Biol. 171, 61–93.
[4] Carrington, J.C., Morris, T.J., Stockley, P.G. and Harrison, S.C. (1987) J. Mol. Biol. 194, 265–276.
[5] Morgounova, E.Yu., Mikhailov, A.M., Kaftanova, A.S., Kulinich, A.V. and Vainshtein, B.K. (1984) Dokl. Acad. Nauk SSSR 276, 119–122 (Russian).
[6] Bricogne, G. (1976) Acta Crystallogr. A32, 832–847.
[7] Acharya, R., Fry, E., Stuart, D., Fox, G., Rowlands, D. and Brown, F. (1989) Nature 327, 709–716.
[8] Love, R.A. and Stroud, R.M. (1986) Protein Eng. 1, 37–46.
[9] Brunger, A.T., Kuriyan, J. and Karplus, M. (1987) Science 235, 458–460.
[10] Ericson, J.W., Rossmann, M.S. (1982) Virology 116, 128–136.
[11] Hogle, J., Kirchhausen, T. and Harrison, S.C. (1983) J. Mol. Biol. 171, 95–100.
[12] Soger, P.K., Stockley, P.S. and Harrison, S.C. (1986) J. Mol. Biol. 191, 639–650.
[13] Kabsch, W. and Sander, C. (1983) Biopolymers 22, 2577–2637.